PSI-PR-96-11
March 1996

The Resolving Power of a Reaction

Frank Cuypers

cuypers@pss058.psi.ch
Paul Scherrer Institute, CH-5232 Villigen PSI, Switzerland

Abstract

We describe a simple method to compute the Cramér-Rao limit of a high energy experiment, *i.e.*, the smallest error with which a parameter can in principle be determined in a reaction. This precision remains a theoretical paradigm since it assumes perfect experimental conditions. Nevertheless, it is shown at hand of an example that for simple processes this asymptotic resolving power can be approached very closely. In all situations, the procedure is at least a useful test of what *could* and what *cannot* be measured by studying a particular reaction.

1 Introduction

It is customary in high energy physics to anticipate experimental results and to determine many years in advance of an experiment how precisely it can measure a parameter. For instance, in the past few years a true industry has been developed to estimate the discovery potential of LEP II. In particular, the reaction $e^+e^- \to W^+W^-$ is a prime candidate for testing anomalous gauge couplings, since it involves the as yet unprobed $WW\gamma$ and WWZ vertices. Typically, one assumes a particular form for these couplings (generally, their standard model prediction) and then proceeds to determine the expected experimental error bounds around this central value.

In general, this procedure depends on four ingredients:

- A theory (*e.g.*, the standard model, its supersymmetric extension, *etc.*) which depends on one or more parameters (couplings, masses, *etc.*). It is the precision with which these parameters can be determined we wish to compute.
- A reaction characterized by its initial and final state (*e.g.*, $e^+e^- \to \mu^+\mu^-$ with or without polarization). This reaction should of course be as sensitive as possible to the values taken by the parameters.
- An observable of this reaction (*e.g.*, the total cross section, asymmetries, *etc.*). It should obviously also depend as much as possible on the parameters.
- A consistent, unbiased and efficient statistical estimator. It is generally chosen to be a least squares or maximum likelihood estimator, which are both equivalent and optimal in the asymptotic limit.

The issue we wish to address here is how to optimize the last two of these four items. For this we shall assume a perfect experiment with no other errors than statistical ones. We shall introduce a *theoretical* observable and a statistical estimator, which yield the smallest possible error on the parameters that can be obtained with a given amount of data. This theoretical limit is nothing else but the Cramér-Rao minimum variance bound [1]. It clearly defines a boundary between what in principle can be achieved and what certainly cannot be achieved, by studying a particular reaction. In the experimental practice, of course, it remains the task of the physicist to make use of an observable (or a set of observables) which yields a sensitivity that comes close to this asymptotic resolving power.

In the next Section we define the χ^2_∞ estimator, which computes the Cramér-Rao limit of the error in the determination of a parameter. In Section 3, we use this criterion to derive limits for an electric dipole moment of the electron in a high energy Møller scattering experiment. Because this reaction is particularly simple it allows the derivation of analytical formulae which nicely exhibit some

general features of the procedure. In Section 4, we consider a similar analysis in Compton scattering. This example will display how realistic a goal the result of the χ^2_∞ estimator can be when the phase space is larger. Finally, we recapitulate in the Conclusion the aim and the domain of applicability of this estimator.

2 The Cramér-Rao Limit

Let us consider a generic high-energy scattering experiment and a theory which by assumption is the correct one. For simplicity we concentrate here on the determination of a single parameter ρ of this theory. It is straightforward to extend all results to follow to the case where several parameters are involved. The true value of the parameter is $\tilde{\rho}$.

We wish to determine the range of values of ρ which would be indistinguishable from $\tilde{\rho}$ when a particular measurement is performed. For example, one could compare the total predicted rates $n(\tilde{\rho})$ and $n(\rho)$. The values of ρ for which

$$n(\tilde{\rho}) - \chi_1 \Delta n(\tilde{\rho}) \quad < \quad n(\rho) \quad < \quad n(\tilde{\rho}) + \chi_1 \Delta n(\tilde{\rho}) \tag{1}$$

cannot be distinguished from $\tilde{\rho}$ to better than χ_1 standard deviations. The average numbers of events n are computed by integrating the differential cross sections over the final state phase space Ω which can be explored by the experiment:

$$n = \mathcal{L}\sigma = \mathcal{L} \int \mathrm{d}\Omega \; \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} \; , \tag{2}$$

where $\mathcal{L}$ is the time integrated luminosity. If systematic errors can be neglected the numbers of events are distributed according to Poisson statistics, and the standard deviation in Eq. (1) is given by

$$\Delta n = \sqrt{n} \; . \tag{3}$$

In order to allow an easy generalization, we can rewrite Eqs (1,3) as a least squares estimator

$$\chi^2_1 = \left(\frac{n(\rho) - n(\tilde{\rho})}{\Delta n(\tilde{\rho})} \right)^2$$

$$= \mathcal{L} \, \frac{\left[\int \mathrm{d}\Omega \, \left(\frac{\mathrm{d}\sigma(\rho)}{\mathrm{d}\Omega} - \frac{\mathrm{d}\sigma(\tilde{\rho})}{\mathrm{d}\Omega} \right) \right]^2}{\int \mathrm{d}\Omega \, \frac{\mathrm{d}\sigma(\tilde{\rho})}{\mathrm{d}\Omega}} \; . \tag{4}$$

The probability that a measurement of ρ deviates from $\tilde{\rho}$ is quantified by χ^2_1: the computed interval of ρ for which χ^2_1 is less than a certain number (say 2.71) will contain a measured value of ρ with the corresponding confidence level (here 90%). The size of this interval is the precision with which the parameter can be determined by measuring the total cross section.

The extent of this error band around $\tilde{\rho}$ depends of course on the value of $\tilde{\rho}$. If experimental data is available, $\tilde{\rho}$ is taken to be the best fit of ρ to this data. In the absence of actual data [1], though, the value of $\tilde{\rho}$ is the result of an educated guess or a theoretical bias, typically, the standard model expectation.

Up to now only a very small portion of the available information has been used. Indeed, it might well be that two very different values of ρ yield the same number of events. Still, these events might have significantly different topologies. Upon integrating over the whole phase space in Eq. (4), these differences are completely washed out. Striking examples of this phenomenon have been discussed in Refs [2].

Clearly, it would be advantageous to include at least some of the information contained in the event shape. This is usually done by considering asymmetries or by dividing the phase space into a certain number N of intervals of one or several kinematical variables $\Delta\Omega_i$ $(i = 1 \ldots N)$. The previous least squares estimator can then be applied separately to each bin in these kinematical variables:

$$\chi^2_N = \sum_{i=1}^{N} \left(\frac{n_i(\rho) - n_i(\tilde{\rho})}{\Delta n_i(\tilde{\rho})} \right)^2 \; , \tag{5}$$

where the index i denotes a particular phase space bin and N is the total number of bins. This is a standard procedure which can substantially improve the resolving power of an experiment [3]. Indeed, because of the triangle inequality χ^2_N can only grow with the number of bins N and one always has $\chi^2_N \geq \chi^2_1$.

Of course, strictly speaking the quantitative probabilistic interpretation of this analysis is only valid as long as the number of bins is not excessive and each bin contains a certain minimum number of events, typically five. Indeed, a χ^2 distribution is defined to be the weighted sum of the squares of independent

[1] This is the situation we consider from now on.

gaussian distributions. However, if too many too small bins are used, this definition is not obeyed for two reasons:

A The binning of the final state phase space takes place with a certain instrumental error, which introduces some amount of bin-to-bin correlation. The numbers of events in different bins are thus not completely independent.

B The number of events in each bin is in reality distributed according to a Poisson distribution, which assumes only asymptotically a gaussian shape.

Obviously, if the number of bins is taken to be so large that the calculated number of events in some bins is less than one, the whole procedure stops making sense.

Notwithstanding this limitation, let us increase (at least on paper) the number of bins to infinity! In this limit the number of events per bin

$$n_i = \mathcal{L} \int\limits_{\Delta\Omega_i} \mathrm{d}\Omega \, \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} \approx \mathcal{L} \, \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\bigg|_{\Omega_i} \Delta\Omega_i \tag{6}$$

is infinitesimally small and χ^2_N (5) becomes

$$\chi^2_\infty = \mathcal{L} \int \mathrm{d}\Omega \, \frac{\left(\dfrac{\mathrm{d}\sigma(\rho)}{\mathrm{d}\Omega} - \dfrac{\mathrm{d}\sigma(\tilde{\rho})}{\mathrm{d}\Omega} \right)^2}{\dfrac{\mathrm{d}\sigma(\tilde{\rho})}{\mathrm{d}\Omega}} \, . \tag{7}$$

Comparing this with χ^2_1 (4), we see that in essence the square of an integral became the integral of a square. Clearly

$$\chi^2_\infty \geq \chi^2_N \geq \chi^2_1 \, , \tag{8}$$

so χ^2_∞ is the most sensitive estimator of ρ.

Because in some sense we assumed an infinite data sample when taking the limit (6), this is the asymptotic resolution which could also be obtained by the maximum likelihood method. Indeed, defining the probability density

$$p = \frac{1}{\sigma} \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} \, , \tag{9}$$

when ρ is in the neighbourhood of $\tilde{\rho}$, χ^2_∞ (7) can be rewritten in the linear regime [2] as

[2] *i.e.* either if the dependence of $p(\rho)$ on the parameter ρ is linear or if the con-

$$\chi^2_\infty = n\,(\rho-\tilde{\rho})^2 \int \mathrm{d}\Omega\ \frac{1}{p}\left(\frac{\partial p}{\partial \rho}\right)^2\Bigg|_{\tilde{\rho}}$$
$$= (\rho-\tilde{\rho})^2 \left\langle -\frac{\partial^2 \ln L}{\partial \rho^2}\Bigg|_{\tilde{\rho}} \right\rangle\ , \tag{10}$$

which is nothing but the maximum likelihood estimator [1], where

$$L = \prod_{i=1}^{n}\ p(\Omega_i) \tag{11}$$

is the maximum likelihood function.

To see that this is indeed the Cramér-Rao minimum variance bound, we set $\chi^2_\infty = 1$ in Eq. (10). Discretizing again into phase space bins, we obtain for the dispersion of ρ around $\tilde{\rho}$

$$D(\rho)^{-1} = \frac{1}{(\rho-\tilde{\rho})^2}\Bigg|_{\chi^2_\infty=1} = \sum_i \frac{1}{n_i}\left(\frac{\partial n_i}{\partial \rho}\right)^2\ . \tag{12}$$

By definition, n_i is the average number of events in bin i. The observed number of events N_i in this bin is distributed according to Poisson statistics, *i.e.*,

$$p_i = \frac{e^{-n_i} n_i^{N_i}}{N_i!} \tag{13}$$

is the probability to find N_i events in bin i. Assuming there are no bin-to-bin correlations, we have

$$< N_i > \ = \ n_i \tag{14}$$
$$< (N_i - n_i)(N_j - n_j) > \ = \ \delta_{ij} n_i \tag{15}$$

and we can rewrite

$$D(\rho)^{-1} = \sum_{i,j} \left\langle \left(\frac{N_i}{n_i}-1\right)\left(\frac{N_j}{n_j}-1\right)\right\rangle \frac{\partial n_i}{\partial \rho}\frac{\partial n_j}{\partial \rho}$$
$$= \left\langle \left(\sum_i \left(\frac{N_i}{n_i}-1\right)\frac{\partial n_i}{\partial \rho}\right)^2 \right\rangle\ . \tag{16}$$

sidered values of ρ are close enough to $\tilde{\rho}$ to warrant sufficient linearity

This is nothing but the Cramér-Rao minimum variance bound [3]

$$D(\rho)^{-1} = \left\langle \left(\sum_i \frac{\partial \ln p_i}{\partial \rho} \right)^2 \right\rangle \ . \tag{17}$$

To derive this result, we only assumed the absence of bin-to-bin correlations in Eq. (15). No assumption concerning the population of the bins is necessary. Although we used the linear approximation in Eq. (12), Eq. (7) remains valid even when the parameter dependence is far from linear, which is often the case when the luminosity $\mathcal{L}$ is small. In contrast, the relations (10) assume a linear parameter dependence because they are derived from the maximum likelihood covariance matrix.

In the presence of real data the maximum likelihood function (11) can easily be evaluated with all experimental resolutions and efficiencies folded in [4]. The linear approximation is then not any longer necessary since the confidence intervals can be estimated without having recourse to the covariance matrix. In contrast, the χ^2_∞ estimator can of course not be applied experimentally, since it assumes (**A**) the absence of systematical errors and (**B**) sufficient statistics to fill infinitesimal bins. These limitations, however, only emphasize the fact that χ^2_∞ yields the theoretical Cramér-Rao limit of what can be measured by the reaction. In other words, any data analysis of a particular reaction, however clever, cannot yield a more precise determination of a given parameter than the asymptotic accuracy yielded by the χ^2_∞ estimator.

If the systematic errors can be neglected with respect to the statistical error, the Cramér-Rao bound predicted by the χ^2_∞ estimator (7) can be experimentally reached with a maximum likelihood analysis. However, if the systematic errors are large, the question arizes, how close can one come in practice to the theoretical precision given by the χ^2_∞ estimator? There is no general answer to this question and a separate analysis has to be performed for each case. This issue is addressed in the next Section at hand of a simple example.

3 Electric Dipole Moment of the Electron in Møller Scattering

To illustrate how the χ^2_∞ estimator works in practice, let us analyze a particularly simple example. If the electron is a composite particle, its non-elementary nature might reveal itself at energies far below it's binding energy by an electric dipole moment d. This dipole plays now the role of the parameter ρ. The

[3] I am indebted to Sergey Alekhin for pointing out this derivation to me.

electron-photon coupling is then described by the effective lagrangian

$$\mathcal{L} = -ie\bar{\psi}\gamma^{\mu}\psi A_{\mu} - i\frac{d}{2}\bar{\psi}\sigma^{\mu\nu}\gamma_5\psi F_{\mu\nu} \ , \tag{18}$$

where e and d are the electromagnetic charge and electric dipole moment of the electron, $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is the strength of the electromagnetic field A_μ and $\sigma^{\mu\nu} = (\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu)$. The first term in the lagrangian (18) represents the *standard* point-like electron-photon coupling, whereas the second term arizes from new interactions.

The static limit for such an electric dipole of the electron is very tightly constrained by low energy experiments [5]. However, such a dipole term might well assume large values for high momentum transfers [6], if it behaves as a function of the photon virtuality Q^2 as

$$d \sim \frac{Q}{\Lambda^2} \ , \tag{19}$$

where Λ is the scale of new physics.

To probe this electric dipole moment of the electron, let us consider a polarized Møller scattering experiment. It has the virtue of being particularly simple and to allow the description of some important features of the χ^2_∞ estimator with handy analytic formulae. The e^-e^- reaction takes place at lowest order in perturbation via the t- and u-channel exchanges of a photon or a neutral vector boson Z^0. In the absence of transverse polarization the final state phase space is one-dimensional. Neglecting the mass of the electron and terms of $\mathcal{O}(d^4)$, the differential cross section for *left-polarized* electron beams becomes

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\cos\theta} = \frac{e^4}{\pi s}\frac{1}{\sin^4\theta}\left(1 + \frac{d^2 s}{2e^2}\sin^2\theta\cos^2\theta\right) \ , \tag{20}$$

where θ is the polar angle of the emerging electrons and $\sqrt{s}$ is the centre of mass energy. To derive Eq. (20) we have ignored the Z^0 exchange. This approximation doesn't introduce any qualitative change, but has the virtue of keeping the analytic expressions simple. In our numerical calculations the Z^0 is of course taken into account.

Such a Møller scattering experiment will be possible at one of the linear colliders of the next generation (CLIC, JLC, NLC, TESLA,...). To be specific, we concentrate here on the canonical design with a centre of mass energy $\sqrt{s} = 500$ GeV and an integrated luminosity $\mathcal{L} = 10$ fb^{-1}. In practice, also, the scattered electrons can only be observed at a certain angle away from the

beampipe. We therefore impose the angular cut

$$\cos\theta < 1 - \epsilon \ . \tag{21}$$

Of course, the resolving power of this reaction depends on the true value $\tilde{d}$ of the parameter. In Fig. 1 the 90% confidence level error band around $\tilde{d}$ (derived from $\chi^2_1, \chi^2_\infty = 2.71$) is plotted as a function of $\tilde{d}$. Since only $|d|^2$ can be observed in this experiment, the plot extends in the same way in the three other quadrants. For (not too) large values of $\tilde{d}$ the resolution scales like

$$\frac{1}{\tilde{d}}\sqrt{\frac{\chi^2}{\mathcal{L}}} \ .$$

Indeed, the expression for χ^2_1 approaches in the limit of a vanishing cut ϵ

$$\chi^2_1 \simeq \frac{\mathcal{L}s}{\pi}\left(d^2 - \tilde{d}^2\right)^2 2\epsilon \ . \tag{22}$$

The reason why χ^2_1 has no sensitivity when the whole kinematical range is inspected ($\epsilon \to 0$), can be traced back to the fact that the dipole moment induces no singularity along the beampipe, in contrast to the point-like coupling. If small angle electrons are also considered, the standard model background keeps increasing whereas the dipole signal does not improve. The collinear divergence of the standard model cross section is eventually regulated by the mass of the electron. Strictly speaking, thus, χ^2_1 in (22) converges to a *very* small but finite value. For our purposes, though, this effect is of no importance.

The angular cut (21) could be optimized (*cf.* Fig. 3) to maximize χ^2_1 [7]. However, a partition of the angular range into a reasonable number N of bins automatically takes care of this task. For the asymptotic limit we find the approximate result

$$\chi^2_\infty \simeq \frac{\mathcal{L}s}{\pi}\left(d^2 - \tilde{d}^2\right)^2 \frac{(1-\epsilon)^5}{10} \ . \tag{23}$$

This is the theoretical limit which can only be approached from below by any experimental setup.

To study the improvement of χ^2_N with increasing number of bins, let us assume the validity of the standard model, *i.e.*, $\tilde{d} = 0$. This way we test the limit of observability of the electron's electric dipole moment. The deviations from $d = 0$ which can be observed with a certain level of confidence (say again 90%) are the values of d which yield a χ^2 in excess of a given number (here again 2.71). In Figs 2 and 3 the d^4 dependence and the angular cut ϵ behaviours of χ^2_1 and χ^2_∞ can be observed to agree with Eqs (22) and (23).

It also appears from Fig. 4, where χ^2_N is plotted as a function of the number of bins N, that with only 30 bins one comes within 90% of the asymptotic resolution. Because the event rates of this reaction are so large, however, the error is in this case dominated by systematics. Assuming for this very clean experiment a .1% systematic error, the expected results are displayed by the dotted curve in Fig. 4.

4 Conclusions

We have presented a simple χ^2_∞ estimator to evaluate the potential of a reaction for studying parameters. The estimator reveals the highest accuracy this reaction could provide under ideal conditions, for determining the numerical values of these parameters: the Cramér-Rao bound.

This estimator does not make any claim about the precision to be obtained under normal running conditions, except that it can never be better. In practice, however, this limit can be closely approached by a maximum likelihood data analysis, if the systematic errors are not too large.

Since the χ^2_∞ estimator provides a bound on what precision can be achieved by a particular reaction in the best of all cases, it is a safe measure to compute this number before embarking on a more time consuming detailed analysis. It can then be decided whether or not this reaction has at all a chance to compete in precision with others.

Acknowledgement

It is a pleasure to thank Sergey Alekhin, Tim Barklow and Geert Jan van Oldenborgh for their enlightening comments and suggestions, as well as Wolfgang Ochs, Stefan Pokorski and Roberto Pittau for their critical comments on the readability of the manuscript.

References

[1] Particle Data Group, *Phys. Rev. D* **50** (1994) 1276;
W.T. Eadie *et al.*, *Statistical Methods in Experimental Physics*, North Holland, 1971.

[2] D. Choudhury, F. Cuypers, *Phys. Lett. B* **325** (1994) 500 [hep-ph/9312308];
Nucl. Phys. B **429** (1994) 33 [hep-ph/9405212].

[3] D. Choudhury, F. Cuypers, A. Leike, *Phys. Lett. B* **333** (1994) 531 [hep-ph/9404362].

[4] T. Barklow, SLAC-PUB-6618.

[5] K. Abdullah *et al.*, *Phys. Rev. Lett.* **65** (1990) 2347.

[6] F. Hoogeveen, L. Stodolsky, *Phys. Lett. B* **212** (1988) 505.

[7] A.A. Pankov, N. Paver, IC-94-274, UTS-DFT-94-10, Sep 1994 [hep-ph/9409395].

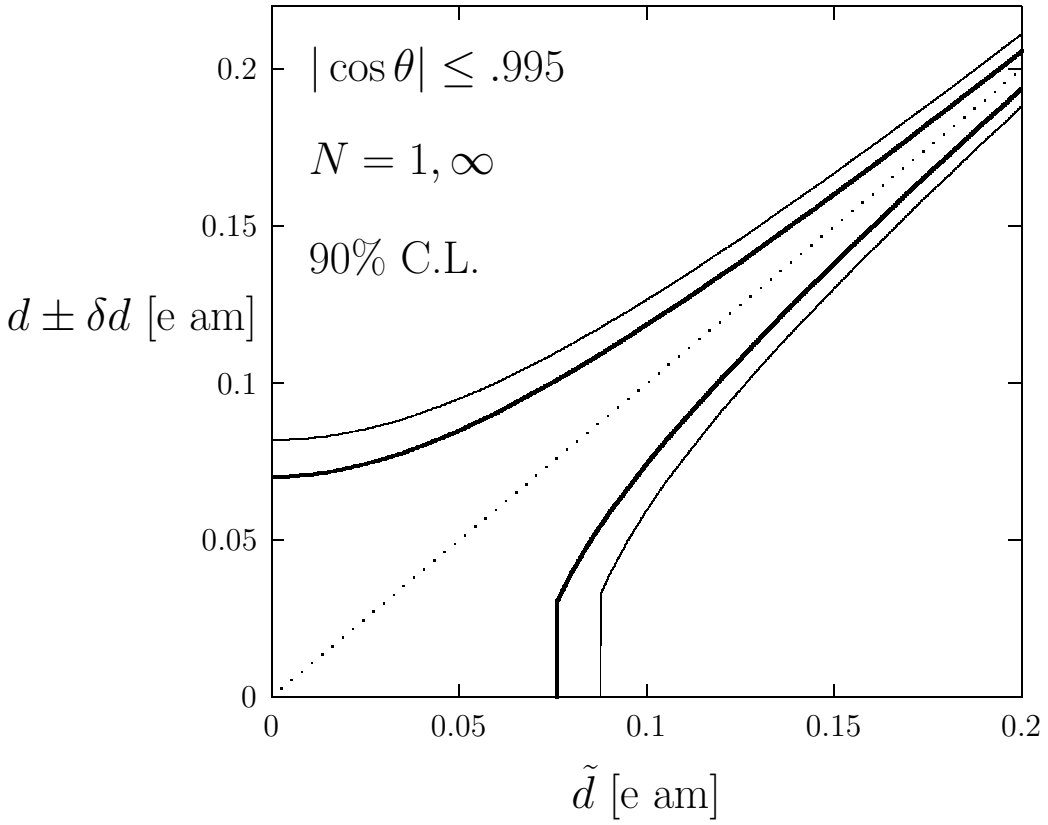

Fig. 1. Dependence of the resolving power on the actual value of the electric dipole moment of the electron $\tilde{d}$. The resolution with one bin (total cross section) and an infinite number of bins (the Cramér-Rao limit) are given by the thinner and thicker curves respectively.

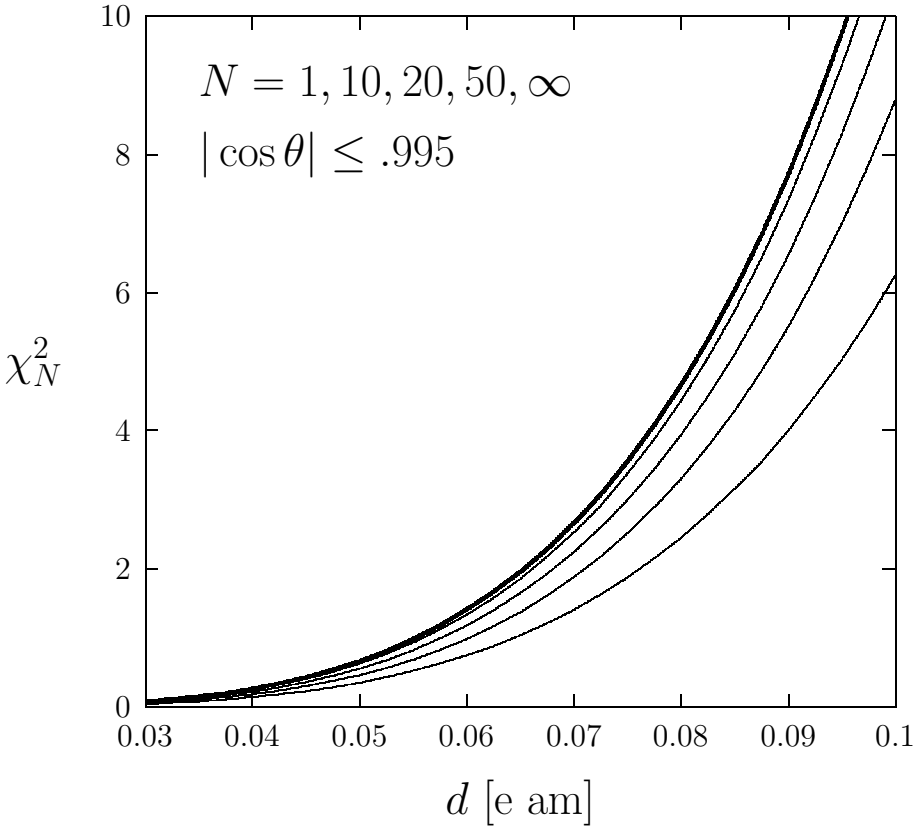

Fig. 2. Dependence of χ^2_N on the electric dipole moment of the electron d.

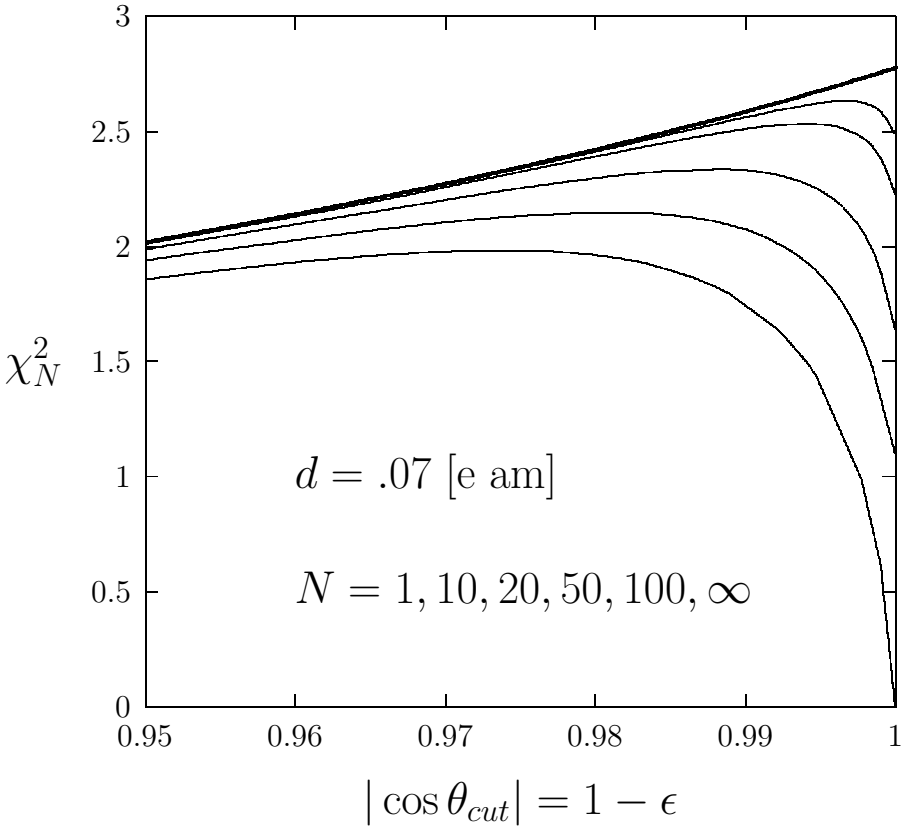

Fig. 3. Dependence of χ^2_N on the angular cut Eq. (21).

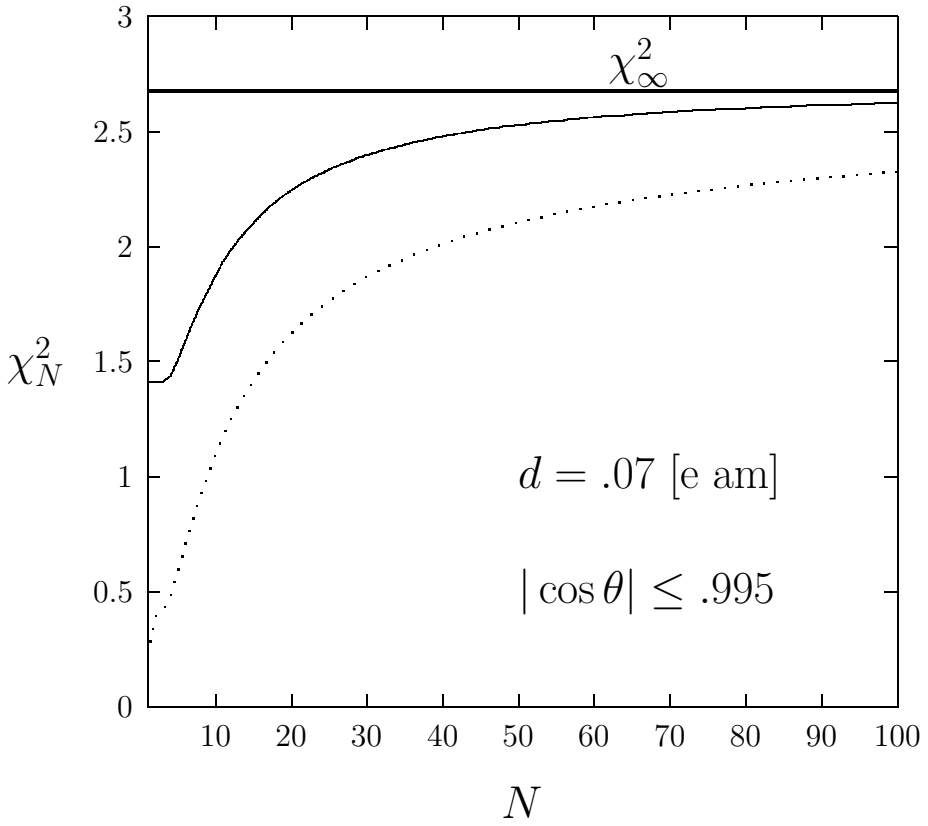

Fig. 4. Dependence of χ^2_N on the number of bins N. The effect of systematic errors is shown by the dotted curve.